

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:52:23 ; Search time 170.72 Seconds  
(without alignments)  
264.889 Million cell updates/sec

Title: US-09-331-631A-3

Perfect score: 3532

Sequence: 1 MAINTSNLCSLFLFLSLFL.....SSRSTKQOQPLVSLDFVGF 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_66:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1191.5	33.7	810	2 T44430	protein PV100 [imp
2	1142	32.3	566	2 S22477	vicilin precursor
3	1092	30.9	605	2 S06398	alpha-globulin typ
4	1083.5	30.7	588	1 FMCNAB	alpha-globulin B p
5	1022.5	28.9	509	2 S08059	globulin-1S, GluB1
6	902	25.5	573	2 A53234	vicilin-like stora
7	895	25.3	582	2 B53234	globulin-1S, GluB1
8	875	24.8	637	2 S35221	globulin-1S, GluB1
9	872.5	24.7	639	2 B24810	beta-conglycinin a
10	859.5	24.3	605	2 S20007	beta-conglycinin a
11	855.5	24.2	605	1 FMS7BA	beta-conglycinin a
12	832.5	23.6	571	2 S00566	convicillin precurs
13	798	22.6	524	2 J01730	62K sucrose-bindin
14	770.5	21.8	483	2 T06459	vicilin-like stora
15	765	21.7	448	1 S24756	vicilin-like stora
16	760.5	21.5	433	2 S00567	vicilin precursor
17	760.5	21.5	459	2 S08505	vicilin precursor
18	749.5	21.2	438	2 S35757	vicilin - garden p
19	748	21.2	439	1 FMS7CB	beta-conglycinin b
20	727	20.6	450	2 S15675	globulin-2 precurs
21	715.5	20.3	463	2 A27288	vicilin precursor
22	711.5	20.1	463	2 S06309	vicilin precursor
23	704.5	19.9	410	1 FWPWB	vicilin precursor
24	689.5	19.5	445	2 J02264	canavalin B - jack b
25	682.5	19.3	445	2 S00281	canavalin - sword
26	568	16.1	436	2 A23498	phaseolin type alp
27	565	16.0	430	2 S10156	alpha-phaseolin pr
28	557	15.8	421	2 A24810	phaseolin beta cha
29	552.5	15.6	414	2 S60190	vicilin - Zamia fu

30	534.5	15.1	540	2 S21825	vicilin-like stora
31	414	11.7	275	1 FWPWA	vicilin A precursor
32	388.5	11.0	236	2 T01662	globulin-1 - maize
33	377	10.7	499	2 T10443	probable major pro
34	373	10.6	407	2 T02258	globulin - maize
35	367.5	10.4	296	2 T06572	convicillin precurs
36	362	10.2	196	2 S02281	beta-conglycinin a
37	355	10.1	218	2 S16334	7S storage protein
38	342	9.7	240	2 T06399	hypothetical prote
39	333	9.4	481	2 T14300	trichohyalin - hum
40	226	6.4	1898	1 A45973	globulin 2 precurs
41	225.5	6.4	136	2 T02870	trichohyalin - rab
42	222	6.3	1407	1 S28589	phaseolin - kidney
43	214.5	6.1	184	1 FSPB	prunin 1 precursor
44	209	5.9	551	2 S51941	trif protein - sil
45	206.5	5.8	1390	2 T14004	

#### ALIGNMENTS

RESULT	1	protein PV100 [imported] - winter squash
T44430		C:Species: Cucurbita maxima (winter squash)
		C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #extl_change 21-Jul-2000
		C:Accession: T44430
		R:Yamada, K.; Shimada, T.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I.
		J. Biol. Chem. 274, 2563-2570, 1999
		A>Title: Multiple functional proteins are produced by cleaving Asn-Gln bonds of a sin
		A:Reference number: 222767; MUID:99107919
		A:Accession: T44430
		A>Status: preliminary: translated from GB/EMBL/DBJ
		A:Molecule type: mRNA
		A:Residues: 1-810 <FAM>
		A:Cross-References: EMBL:AB019195; NID:g3808061; PIDN:BAA34056.1; PID:g3808062
Query Match	33.7%	Score 1191.5; DB 2; Length 810;
Best local similarity	34.3%	Pred. No. 16-66;
Matches 284;	Conservative 144;	Mismatches 218; Indels 183; Gaps 23;
OY	1	MAINTSNLCSLFLFLSLFLSTVSLAE-----SEP- 31
DB	1	MALSKVKLRLLAFTLFLACLISVGLDGKESLSGAGVDHOGVNRCELKGNVDEFA 60
OY	32	-----DRQYECKRQCMQLETS-GQMRCSVQCCKREE----- 65
DB	61	ACKKGGGVNORGSPRAVEYVCRLRCVAERGVQQRKCEQVCEERLRERQGRGDEV 120
OY	66	---DIMSRY-----DNDDPQTDQOQCRKROE 93
DB	121	ERRDPMEERQQRREHERERERERERERERERERERERERERERERERERERER 178
OY	94	SGPRQOQYOCRKCELEEEVEYNRQDPQOQYEQOCERQARHETPRHMQTCQRCERR 153
DB	179	RRRREQER 232
OY	154	Y-----EKE-----KKQKRYEQQREDEKLEEMKEDKRP-----QOREYE 195
DB	233	RNGRDEDEENQDPDPRRQKRRQEQERERERERERERERERERERERERERER 292
OY	196	DCRRRCQOE-----PROOYQOCRCRCROHGRGG-- 227
DB	293	ERRRRERQER 352
OY	228	-----DLNPORGSGRYEE-----GEKQSDNPYYFEDRSLSFTRTEEGHISVL 273
DB	353	EQSRDEDERRRHRRQGGSRVNOVAIRREDEQSNPPYFQGRQSNVRSDEGHWRVL 412
OY	274	ENYGGSKLLRALKNRLVILEANPNAPVLPHTLDDALLVYGGGALKMHRNR--- 330
DB	413	ERSESESLKIGIKNORLALLEARPHTFIVPHILADECVLLVVRGRATITVVOEKRETR 472

QY	331	-ESYLEGGDQVITPAGTTFYILINRDQNEBLIAEFLTOTISPGOYKEFPFGQNPDEY	389
Db	473	KESTYVESGDQWTTIPAGTTLTIANE-NEDLOIVLQVPYNNPGFEKDYLSAGGGS-QAY	530
QY	360	LSTFKEILLEALNTOAERLKGVLGQRE--GVIIISASQEOITRELTRODDESRRHHIRRG	447
Db	531	YSVFENDDLEALNLTIPROKLERIFKORRERGKIVRASQEOIRALSO-----RATSVRKG	585
QY	448	GESSGCPNLFKRPPLYSNKYGOAVEYKPEDYROLQDMDVSFIANITOGSMGCPFMTNR	507
Db	566	SRGVNAPITKLEBQTFYVNNQYQGMFEACPDPERPOLRRDVTATSVYDIOGGIMVHFMSR	645
QY	508	STKYVYVASGEADVMACPHLS-----GRHGGRGGKRRHEEE---DVHTEQYKARLS	557
Db	646	ATWVVFVEBGAGSFEMACPHTIOSQOMQRGRREERHMRREDEEREERSGRGERVARGRLS	705
QY	558	KREALIVVGPVYVSSGGENILTFAGFINAQNHNHEPLAGRENVYQOIEPQOMELAF	617
Db	706	EGGVLTIVPAGHIALIMASPENELRLVGFGINAMENHRRPLAGRE-NINNELDREKLELAF	764
QY	618	AAPKREVELLNSQDSEITFFPPQPRHOHQOOSKSTKYOOOPLYSLDPVGE	666
Db	765	NVEGQAOEITFRSQRESSEFTGEPGGRRRST---ERSPLTILKLAJY	809

RESULT 2

S22477

vicillin precursor - cacao

C:Species: Theobroma cacao (cacao)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 17-Mar-1999

C:Accession: S22477; S22478; S18105; S22050

R:McHenry, L.; Fritz, P.J.

Plant MOL. Biol. 18, 1173-1176, 1992

A:Title: Comparison of the structure and nucleotide sequence of vicillin genes of cocoa

A:Reference number: S22477; MUID:92288309

A:Accession: S22477

A:Molecule type: DNA

A:Residues: 1-566 <MC2>

A:Cross-references: EMBL:X62625

A:Accession: S22478

A:Molecule type: mRNA

A:Residues: 1-452 <MC2>

A:Cross-references: EMBL:X62626

C:Genetics:

A:Introns: 211/1; 269/3; 296/3; 391/3; 502/1

C:Superfamily: glycinin

C:Keywords: seed; storage protein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-566/Product: vicillin #status predicted <MAT>

```

Query Match      32.3% Score 1142: DB 2: Length 566;
Best Local Similarity 41.1% Pred. No. 1.2e-63;
Matches 237; Conservative 106; Mismatches 176; Indels 58; Gaps 13;

QY 109 ICEEEEEYNR--ORDPOOYEOCQERCORHETEPHRMOTOCORCERYEKEKRRQOKRY 165
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 22 LCSVSASAGFRKQYRDPRPOOYEOCCORCESEATEEREQEOECORCER-----EY 70

QY 166 EEOREDEBEKEEEMKEEDNKRDPQOREYEDCRRCEQDE--PPOOYOCORCRCEQOKH 223
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 71 KEOROOOEEL-----ORYOOCQGGCQDQOQGRDQOKHCEWOYEQ 116

QY 224 GRGDDLINPONGSGRGYECEGKOSDNPPYFDE-RSLSTRTEGHTSVLENFGRSKL 282
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 117 ER-GEHENYHNHKKNRSEEGQGRNPPYFRPKRRSFQTRRDEGKNTLQRAENSPP 175

QY 283 LRALKNRIYLLKNPNAFYLPHTLDDATLLVYGGKALKMHRDNRESNINLECGDIYR 342
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 176 LKGINDRILAFMEANPNFTFLPHDCDAEAYEVYNGKSTLFEVHNESKSNIVORGTVS 235

QY 343 IPACTTYYLINRDNNERHLIAKFDOTISTPGOYKEFFPAGGONDEPYLSTYSKEILLEAL 402

```

[illegible]

```

RESULT      3
S06398
alpha-globulin type A precursor - upland cotton
N:Alternate names: seed storage protein
C:Species: Gossypium hirsutum (upland cotton)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
C:Accession: S06398
R:Chian, C.A.; Borriolo, K.; Kamalay, J.A.; Dure III, L.
Plant Mol. Biol. 9, 533-546, 1987
A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX
A:Reference number: S06398
A:Accession: S06398
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-605 <CHL>
C:Superfamily: glycinin
C:1-24/domain: signal sequence #status predicted <SIG>
C:25-605/Product: alpha-globulin type A #status predicted <MAT>

```

	Query Match	30.9%	Score 1092;	DB 2;	length 605;
	Best Local Similarity	36.2%;	Pred. No. 1.7e-60;		
	Matches 249;	Conservative 118;	Mismatches 210;	Indels 110;	Gaps 17;
QY	1	MAINNSNLCISLLEFLSLFTSTVSLAE-----SEFD-RQLEYECKRCQMLNLSGQMR	53		
Db	1	MYRNKSVFVLEFL--SLFLSFGLLCSANDPGRSRSEDDPQGRYEDCRRKN-OLETRGQRE	57		
QY	54	RCVSCDCKRFREEDIMWSYANDODDQOTQOQOQRCRCROQESPRQOQYQCRBKCEI	113		
Db	58	Q-DKCEBRSETQLNEEQRQDEQDORRYQDORQHCQOEERLPR-----HCEQSCRQO	109		
QY	114	EEYNKORPQOQYEQSCQRCQRPETPRNMQSCQRCSERYEKEKQKQRYEEOQREDE	173		
Db	110	YEKQOQOQPRDKQFKKQOQRCQOQRCQOQREPRKQOQYACRCQOYQEDWIKGR	160		
QY	174	EKEERMKCEDNKRDPQOREYEDCRRRCQEORQRYOQCRRCQREQORNGRGDLINPQ	233		
Db	161	NKMRREEEES-----	171		
QY	234	RGSGRGEEGEEKOSDNFYEDERSLSTRFTEECHLSYLENFYGRSKLLRALKNRYLV	293		
Db	172	-----DEGEQOQNNNYTFPIHRSRQFERFEGREHGNFRQLQCFADKNNHLKRI	224		
QY	294	LEANPNAFVLPFLTHDADAILLVTGSGALAKMITHRDNRESYNLECGDVIIRIPAGTTFYLN	353		
Db	225	LEANNTFVLPHHCHCAEKIIVYVITNGRGIVTFYTHNKESYINVPQVVVIRIPAGSTVIYLAN	284		
QY	354	RDNNERLHIAKFLQITSPGOYKEFFPPAGQONDEPYLSTFSKEILAEALNTQAEIRGLV	413		

[illegible]

RESULT 4

FCNMB  
alpha-globulin B precursor (clone C72) - upland cotton  
N:Alternate names: seed storage protein; vicilin precursor  
C:Species: *Gossypium hirsutum* (upland cotton)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999  
C:Accession: A30838; S06911  
R:Chan, C.A.; Pyle, J.B.; Legocki, A.B.; Dure III, L.  
Plant Mol. Biol. 7, 475-489, 1986  
A:Title: Developmental biochemistry of cottonseed embryogenesis and germination XVIII cl  
A:Reference number: A30838  
A:Accession: A30838

A:Molecule type: mRNA  
A:Residues: 1-588 <CH2>  
A:Cross-references: GH:M16891; NID:g167374; PIDN:AAA33071.1; PID:g167375  
A:Experimental source: Var. Coker 201  
R:Chan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.  
Plant Mol. Biol. 9, 533-546, 1987  
A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Se  
A:Reference number: S06398  
A:Accession: S06911  
A:Status: not compared with conceptual translation

A:Molecule type: DNA  
A:Residues: 1-81 <CH2>  
A:Comment: This is a seed storage protein.  
C:Superfamily: glycinin  
C:Keywords: glycoprotein; seed; storage protein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-386/Product: alpha-globulin storage protein #status predicted <MAT>  
F:417/Blinding site: carboxydrate (Asn) (covalent) #status predicted

Query Match	30.7%;	Score 1083.5;	DB 1;	Length 588;
Best Local Similarity	38.2%;	Pred. No. 5.5e-60;		
Matches 229;	Conservative 114;	Mismatches 182;	Indels 75;	Gaps 13;

OY	76	DDPOTDQOQORRCRODESGPRQOYOQRCKELCEELEYNQR--DPQOYECCOEK	133
		:      :     :   :	
Dd	35	DDPKRYEDCRRKEWMDTRQCEOOOCESCKSOYGKDKDOHNRBDPPRREYECOC	94
OY	134	ORHTERRHMJOTCOORERREYKKRKQORYEQOHEDEBEKYEHRKLEDNNKRDRQORE	193
		: :        :  :  :  :  :	
Dd	95	R--QOERFORQOCOQLKCRFBEOO-----SQRO	123
OY	194	YEDRCRRCEODE--PROYOCCRRCFEODRONRGDLINPOR-----CGSRVEEGEE	245
		:::::      : : :      : : :	
Dd	124	FQEQQCHHOQENRPKKQOCVAKRCGRARYE-----NPMKREERELEEETEBEGD	175
OY	246	KOSDNPYFDERSLSTRTEEGHISYLENFYGRSKLLRALKNYRVULLEANNAFLPT	305

```

Db      176  EOSHNPFEHFRHSQSRPREHGHFNRLQRFASHPLLRIGINEFRLLSEANPNFVLRH 235
Oy      306  HLDADAILLVTVGGALAKMIRDRRESYVINECGVITRTPAGTFFYLINDNNRHLIAKF 365
Db      236  HODAEKTYLVNNGSTLFTLTHEKKEYNVPYGVYKVPAGSVIYLANDNNKEKLIYAL 295
Oy      366  LQIISTPGOYKEFPFAGQNPEDPLSTFSKEILEALNTQAEKRLGVLG-----QORE 418
Db      296  HRPVNNPQOFEFFPPASQSRQSTLRAFSEKELLEPAFNTRSDLDLFGROSHRRQOQO 355
Oy      419  GVIIISASQEOIRRELTDRDSESRMHIRRGESSRGYPVFNKRPVLSNRYGAYEYKPED 478
Db      356  GMRKASQEOIRALSQENTSPR-----EKSE--RFAFNLLYRPRRSNONGRYEACPRE 409
Oy      479  YROLODMVSYFANITTOGSMGCFPFNTSTKYVVVYVAASGADVEMACPILSGRHGRGG 538
Db      410  FROSLDINVTWSALQNGQSIIFYPHYNSKATFVPLVNEGNGYEVNYSPIILRQSSPEEE 469
Oy      539  KRHEEEDV-----HYEQYAKRLSKREALVVPVGHVYVSSGNEMLLFAFG-----I 587
Db      470  EOOEOEOEEDERRSGOYKRIKRSOLSRDIEVVPANPVPYFVASOONLMLTGTGGLYNOI 529
Oy      588  NAOINNHEMFLAGRERNTLOQIEPQAMLAAPARKVEDELFNSQDSIFFRPGRHOQOS 647
Db      530  NPDHQRLEFVAGKLNHY--RQWDSQAEKLAVGSRLVDEFNPNPDESVFVS--RORORAS 587

```

```

result      5
S08059
alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)
N:Alternate names: seed storage protein
C:Species: Gossypium hirsutum (upland cotton)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
C:Accession: S08059
R:Chian, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
Plant Mol. Biol. 9, 533-546, 1987
A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.
A:Reference number: S06398
A:Accession: S08059
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1*509 <CHL>
C:Superfamily: glycinin

```

Query Match	28.9%;	Score 1022.5;	DB 2;	Length 509;
Best Local Similarity	38.8%;	Pred. No. 2.8e-56;		
Matches 217;	Conservative 105;	Mismatches 153;	Indels 85;	Gaps 13;

QY	121	DPDOYECOSGCRNHEPRNMTCCOORGRRTYKFKRKOQKRYEDQREDDEKYEERM	180
Db	1	DPORRETCOOCBCR--QOEERKOOPCOOCRLKREDEOEO-----	38
QY	181	KEEDNRDPQOREYEDCRRRCOEQ--PRDYOCORRREOOROHGRGDLINORG--	235
Db	39	-----SQRFOEQCHQOEOREPKKQOCVCRECKEYQ-----NWRGERE	81
QY	236	---GSGRYEEEEKOSDPRYFYFDRSLSTRRTREBGHSLYENLYGSKLLRALKNRLV	29
Db	82	EEABEETEBEEOESHPPFHHRFSQSRREEHGNRYLQRPASRHPILIRGNERLS	141
QY	293	LLLEANPAFLVPLTHADAILLVGSGKALKIMIRDRNRESYNLCCGVIRIIPAGTTFYLI	352
Db	142	ILLEANPFLVPLRHCDAREKIVLYTNGKOTLFLTHENKESYNIVGVVWVPGSVYYLA	201
QY	353	NRDNNRHLIAKFLQTSITPGOYKEPFRPAGQONEPRLTFSKILLEALANTQAEKRGV	412
Db	202	NQDNKKELTIVLHRPVANNPQOFEEFPAGQORQOSYLRAASRRLIPLAFRTSEQDEL	261
QY	413	LG-----QOREGVIIISASODQIRELIRDDSESRMHIRKGCSSGCPNLFKRPLYS	465
Db	262	EGGRSRRRQOGCMERKRASEOIRALSOEATSPR---EKSGE--RFAFLLISOPRYS	315

QY	466	NKYGAEVKKPEPDYVQLODMQVSYFANTIQSSMMGPFENFTNSTKVVVYVASEADYEMAC	525
Db	316	NONGFEFFECCEPEFFQLODRINTVYALADNOSGIFVPHYNSATFVILTEGNGYAEWVS	375
QY	526	PHLSGRHGRGRCGRKHEEBEDV-----HYEOKVRLSKREAIIVPVGHPVFE	572
Db	376	PHLP-----ROSYEEFEFEDEBEEOGEERKRSQYKRIRSRILSGDIFVYPANFPYFE	430
QY	573	VSSGNENLLFAFG-----INANNHENEFLAGREBNVLOQIEPQAMELAFAPRKEVEEL	627
Db	431	VASQNONLRMTGCGFLAYNONINPDHQRITFVAGKINHV--ROMDSQAKELAFGVSSRLVDEI	489
QY	628	FNSODESIFFCGRQHOOS	647
Db	430	FNSNPOESYFWS--RQRORAS	508

```

RESULT      6
A53234      globulin-1S, GIBIS - maize
C:Species:  Zea mays (maize)
C:Date:      02-May-1994 #sequence_revision 18-Nov-1994 #text_change 11-Jan-2000
C:Accession: A53234; A43642
R:Belanger, F.C.; Kriz, A.L.
Genetics 129, 863-872, 1991
A:Title:     Molecular basis for allelic polymorphism of the maize Globulin-1 gene
A:Reference number: A53234; MUID:92090707
A:Accession:  A53234
A:Status:     preliminary
A:Molecule type: DNA
A:Residues:   1-573 <BEL>
A:Experimental source: Inbred line Va 26
A:Note:       sequence extracted from NCBI backbone (NCBIN:71280, NCBIPI:71284)
R:Kriz, A.L.
Biochem. Genet. 27, 239-251, 1989
A:Title:     Characterization of embryo globulins encoded by the maize G1b genes.
A:Reference number: A43642; MUID:89374022
A:Accession:  A43642
A:Status:     preliminary
A:Molecule type: protein
A:Residues:   87-100 <KRI>
A:Superfamily: glycinin

```

Query Match	25.58	Score 902	DB 2	Length 573
Best Local Similarity	39.98	Pred. No. 9.9e-49		
Matches 203; Conservative	84	Mismatches 180	Indels 42	Gaps 11

[illegible]

```

OY 523 MACPILSRHG----GRGCGKHREDEVHYEQ-----VKARLSKREALVVPVPGH 569
Db 379 IYCPHQSGGSEBSEBKEKGRSEBEEBESSEDEGEVQGYHTIRARLSPTATVVPVAGPH 438
OY 570 VYFVSSGJENLLIFAFGJNMOINNIENLAGERVLQOIEPOMELAFAPKREVEYELFN 629
Db 439 FVAVASRSLNIQICFEVHADRNKEVFLAGND-VYLOKLDPRVAKULSFASKAEDEVVLQ 497
OY 630 SODESIFFFGPRQ--HQQGSNSTKQOO 655
Db 498 SNREGFLPGKESGJHEERQEEBEE 526

```

```

RESULT      7
B53234
vicilin-like storage protein G1b1-L, embryo - maize
N.Alternate names: globulin-1L
C.Species: Zea mays (maize)
C.Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C.Accession: B53234; S21824
R.Belanger, F.C.; Krlz, A.L.
Genetics 129, 863-872, 1991
A.Title: Molecular basis for allelic polymorphism of the maize globulin-1 gene
A.Reference number: A53234; MUID:92090707
A.Accession: B53234
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-582 <BRL>
A.Cross-references: EMBL:X59083; NID:g22283; PIDN:CAA41809.1; PID:g22284
A.Experimental source: Inbred line W64A6
A.Note: sequence extracted from NCBI backbone (NCBIP:71285)
C.Genetics:
A.Gene: G1b1-L
A.Introns: 167/1; 225/3; 252/3; 349/3
C.Superfamily: glycinin

```

Query Match	25.3%	Score 895	DB 2	Length 582
Best Local Similarity	39.6%	Pred. No. 2.7e-48		
Matches 201; Conservative	83;	Mismatches 183;	Indels 40;	Gaps 11

[illegible]



Db 495 VLVINEGANIELV-----GIKEOQOQOEOPLEVRKRYRAELSEODIFVIPAGY 546  
QY 569 PVFVSSGNENLLFPAGINANNHENPLAGRENVLOOIEPOAMELAPARKVEEELF 628  
Db 547 PVWV--NATSDLNFPFAGINANNONRFLAGSDNVISQIPSOVELAPFRSAKDIENLI 604  
QY 629 NSODESIFPPGPRHOQOSSRSTKQOOPVLSIL 661  
Db 605 KQSSESFYVDAQOQOEKGNKGRK--GPLSIL 635  
RESULT 10  
S20007  
beta-conglycinin alpha chain precursor - soybean  
C:Species: Glycine max (soybean)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
C:Accession: S20007  
R:Jellevre, J.M.; Dickinson, C.D.; Dickinson, L.A.; Nielsen, N.C.  
Plant Mol. Biol. 18, 259-274, 1992  
A:Title: Synthesis and assembly of soybean beta-conglycinin in vitro.  
A:Reference number: S20007; MUID:92119248  
A:Accession: S20007  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-605 <DEL>  
C:Superfamily: glycinin

Query Match 24.3%; Score 859.5; DB 2; Length 605;  
Best Local Similarity 31.6%; Pred. No. 4,6e-46;  
Matches 195; Conservative 143; Mismatches 210; Indels 69; Gaps 15;  
QY 72 YDMDDPOTDCOQORRCROESGPRQOQYCORRC-----KEICEEE--EYNQORDP 122  
Db 27 YWKEKNPKNH--KCLQSCNSERSYRNQA-CHARNILKYEKECEGELPRPRRQPH 83  
QY 123 QOOYEOQRCORHETPRHMOTCOQRCERRYEKEKROOKRYEEOQREDEEYERME 182  
Db 84 ERPRQOPGKEEDEDQPRIPPRPQOPROEHEEOREDEWPKREKKEGKESEEDD 143  
QY 183 EDKRPQOREYEDRCRRQOEPRQOYOCQRCROQROHNGSGDLINQRGSSRYEE 242  
Db 144 EDEQDERQPRP--RPRHQKEERKEDEED--EEOQRS-----EE 181  
QY 243 GEEKQ-----SDNPVYFEDRSLSRTRETEGHTSVLENFYGRSKLRLAKNRYLVLEAN 297  
Db 182 SEDSELRHRKKNKPRFLGSRNRFETLFKNQYGRIVLQRFNRQSRPOLQNLRDYILEFNSK 241  
QY 298 PNAFVLPHTLDADAILLVGGRGALMIRHNDRESYNLECGDYIRIPAGTTFYLINRDN 357  
Db 242 PNTLLPNNHADADYLVILNGTALLSLVNNDDSDYRLQSGDALRVPSGTTYVVPDNN 301  
QY 358 ERLHIKFLQITSTPGQYKEFPFAGQONPRPYLSTESKELLEALMTQMERLGVL---- 413  
Db 302 ENLRLLTLAIPVKKPGFESFELSTEDAOOSYLQGSFNRNLEASDYTKFEELIKVLFESRE 361  
QY 414 -----GOOR--EGVILISASOEQIRELTRDSESRMHIRGSGSSGPNLFFKRPRLYSN 466  
Db 362 EGOQOGEORLOESVYIEISKEQIRALSKAKSSSKRTI-----SSEKPPNLSRDPYISN 417  
QY 467 KYGQAVEPREDYROLQDMQVSVFANITQSGMGPFFTRSTKVVVVAAGEADVEMACP 526  
Db 418 KICKFPEITPEKPNQRLDIFLSYDMEGALLPHFNSKAIIVLIVINEGDANIELV-- 475  
QY 527 HLSGRNG--RRGKRHNEEEDVHYEQVAKRLSKRAIYVPGHPRVVPVSSGNENLLFA 564  
Db 476 -----GKEQOQOEQOQOEPRLEVRKRYRAELSEODIFVIPAGYVWV--NATSNLNFA 526  
QY 585 FGINANNHENFLAGRENVLOOIEPOAMELAPARKVEEELFNSODESIFPPGPROH 644  
Db 527 IGINANNONRFLAGSDNVISQIPSOVELAPFGSAQAVAEKLLKQRESYFYVDAQPKK 566  
QY 645 QOSSRSTKQOOPVLSIL 661

Db 587 EDEGNKGRK--GPLSIL 601  
RESULT 11  
FWSYBA  
beta-conglycinin alpha chain precursor - soybean  
C:Species: Glycine max (soybean)  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jul-1999  
C:Accession: S14681; S74124; S06714  
R:Sebastian, F.L.; Farrell, L.B.; Schuler, M.A.; Beachy, R.N.  
Plant Mol. Biol. 15, 197-201, 1990  
A:Title: Complete sequence of a cDNA of alpha subunit of soybean beta-conglycinin.  
A:Reference number: S14681; MUID:91355860  
A:Accession: S14681  
A:Molecule type: mRNA  
A:Residues: 1-605 <SEB>  
A:Cross-references: EMBL:X17698; NID:q18535; PID:CA935691.1; PID:q18536  
R:Shutov, A.D.; Kakhovskaya, I.A.; Bastirgina, A.S.; Bulmaga, V.P.; Horstmann, C.; Mu  
Eur. J. Biochem. 241, 221-228, 1996  
A:Title: Limited proteolysis of beta-conglycinin and glycinin, the 7S and 11S storage  
F:261.517/Binding site: carbohydrate (Asn) (covalent) #status predicted  
A:Accession: S74124; MUID:97054613  
A:Accession: S74124  
A:Molecule type: protein  
A:Residues: 189-196, 'H', 198, 'N', 200, 'X', 202-203, 397-408, 'X', 410, 'X', 412-417, 'X', 419-4  
A:Experimental source: seed  
C:Superfamily: glycinin  
C:Keywords: glycoprotein; seed; storage protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-62/Domain: propeptide #status predicted <PRO>  
F:63-605/Product: beta-conglycinin alpha chain #status predicted <MAT>  
F:261.517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.2%; Score 855.5; DB 1; Length 605;  
Best Local Similarity 31.4%; Pred. No. 8,1e-46;  
Matches 194; Conservative 143; Mismatches 211; Indels 69; Gaps 15;  
QY 72 YDMDDPOTDCOQORRCROESGPRQOQYCORRC-----KEICEEE--EYNQORDP 122  
Db 27 YWKEKNPKNH--KCLQSCNSERSYRNQA-CHARNILKYEKECEGELPRPRRQPH 83  
QY 123 QOOYEOQRCORHETPRHMOTCOQRCERRYEKEKROOKRYEEOQREDEEYERME 182  
Db 84 ERPRQOPGKEEDEDQPRIPPRPQOPROEHEEOREDEWPKREKKEGKESEEDD 143  
QY 183 EDKRPQOREYEDRCRRQOEPRQOYOCQRCROQROHNGSGDLINQRGSSRYEE 242  
Db 144 EDEQDERQPRP--RPRHQKEERNEEDED--EEOQRS-----EE 181  
QY 243 GEEKQ-----SDNPVYFEDRSLSRTRETEGHTSVLENFYGRSKLRLAKNRYLVLEAN 297  
Db 182 SEDSELRHRKKNKPRFLGSRNRFETLFKNQYGRIVLQRFNRQSRPOLQNLRDYILEFNSK 241  
QY 298 PNAFVLPHTLDADAILLVGGRGALMIRHNDRESYNLECGDYIRIPAGTTFYLINRDN 357  
Db 242 PNTLLPNNHADADYLVILNGTALLSLVNNDDSDYRLQSGDALRVPSGTTYVVPDNN 301  
QY 358 ERLHIKFLQITSTPGQYKEFPFAGQONPRPYLSTESKELLEALMTQMERLGVL---- 413  
Db 302 ENLRLLTLAIPVKKPGFESFELSTEDAOOSYLQGSFNRNLEASDYTKFEELIKVLFESRE 361  
QY 414 -----GOOR--EGVILISASOEQIRELTRDSESRMHIRGSGSSGPNLFFKRPRLYSN 466  
Db 362 EGOQOGEORLOESVYIEISKEQIRALSKAKSSSKRTI-----SSEKPPNLSRDPYISN 417  
QY 467 KYGQAVEPREDYROLQDMQVSVFANITQSGMGPFFTRSTKVVVVAAGEADVEMACP 526  
Db 418 KICKFPEITPEKPNQRLDIFLSYDMEGALLPHFNSKAIIVLIVINEGDANIELV-- 475  
QY 527 HLSGRNG--RRGKRHNEEEDVHYEQVAKRLSKRAIYVPGHPRVVPVSSGNENLLFA 564  
Db 476 -----GKEQOQOEQOQOEPRLEVRKRYRAELSEODIFVIPAGYVWV--NATSNLNFA 526



